JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **166**, 404–427 (1992)

Strongly Consistent Estimation of the Order of Stochastic Control Systems (CARMA Model)

E. M. HEMERLY

Instituto Tecnológico de Aeronáutica, CTA-ITA-IEEE, 12225-São José dos Campos, São Paulo, Brazil

AND

M. D. FRAGOSO

Laboratório Nacional de Computação Científica, LNCC/CNPq, Departamento de Pesquisa e Desenvolvimento, Rua Lauro Müller 455, 22290-Rio de Janeiro, Rio de Janeiro, Brazil

Submitted by E. Stanley Lee

Received May 14, 1990

In this paper a variant of the PLC(Predictive Least Squares) criterion for order estimation is combined with an adaptive control strategy and applied to multidimensional CARMA systems. It is shown that with this combination we can estimate, recursively and in a strongly consistent way, both the order and coefficients of the controlled system, while achieving asymptotically optimal cost.

1. INTRODUCTION

The PLS(Predictive Least Squares) criterion for order estimation was recently introduced by Rissanen [14] and its asymptotic properties were analysed by Hemerly and Davis [8], Hannan, McDougall, and Poskitt [7], and Veres [16]. When compared with classical criteria for order selection, like those in Akaike [1], Rissanen [12, 13], Schwarz [15], and Hannan and Quinn [6], the PLS criterion has substancial advantages in terms of recursive computability (see Wax [17] for details).

The aforementioned works were carried out in the time series framework, where some sort of stability and ergodicity of the stochastic processes is assumed. These results can not be directly applied to systems operating under feedback, for these basic assumptions cannot be ensured a priori. The first reference to address the order estimation problem of systems operating under feedback appears to have been Chen and Guo [3]. They introduced the L_n criterion, which is a modification of the

0022-247X/92 $5.00

BIC criterion, to estimate the order of multidimensional ARX systems operating under adaptive control of the LQG type, proposed by Chen and Guo [2]. They proved that this combination provides consistent estimates for both the order and coefficients of the controlled system. In Guo, Chen, and Zhang (1987) this result is generalized to CARMA models by the introduction of a new criterion, the CIC.

Recently Hemerly and Davis [9] showed, for ARX models, that by combining the PLS criterion with a control strategy similar to that in Chen and Guo [3] it is possible to devise an asymptotically optimal adaptive control strategy and simultaneously estimate, recursively and in a strongly consistent way, both the coefficients and the order of the controlled system.

This paper concerns the extension of the results in Hemerly and Davis [9] to systems with correlated noise. To cope with the stronger data correlation in this case, we introduce a variant of the PLC criterion.

2. Problem Formulation and Technical Assumptions

In this paper we consider systems represented by multidimensional CARMA models, with the parameters being identified via the AML(Approximated Maximum Likelihood) technique, whose main advantage when compared with the general RPE(Recursive Prediction Error) consists in not requiring projection into the stability region of certain polynomials (see Ljung and Söderström [11]).

Due to the stronger data correlation in the present set-up, it has not been possible to establish a result similar to that presented in Hemerly and Davis [9]. More precisely, in order to establish the strong consistency of the order estimate we have now to give up, to a certain extent, one of the most appealing characteristics of the PLS criterion, namely its exclusive dependence on the data. In other words, the criterion we consider in this paper lies somewhere between the PLS criterion in Hemerly and Davis [9] and the CIC criterion considered by Guo, Chen, and Zhang [5], where the data dependent part of the criterion passes over it twice. The improvement we provide here consists in showing that results similar to that of Guo, Chen, and Zhang [5] can be established by using a criterion for order determination whose data dependent part can be evaluated recursively.

Let $(\Omega, \mathscr{F}, \mathscr{F}_t)_{t=0,1,\ldots}, \mathbb{P})$ be a filtered probability space, $\{w(t)\}$ a martingale difference process with respect to $\{\mathscr{F}_t\}$, and consider the l-input, m-output stochastic control system described by

$$\begin{aligned} y(t+1) &= A_1 y(t) + \cdots + A_{p_0} y(t-p_0+1) + B_1 u(t) + \cdots \\ &\quad + B_{q_0} u(t-q_0+1) + w(t+1) + C_1 w(t) + \cdots \\ &\quad + C_{r_0} w(t-r_0+1), \qquad y(t)=u(t)=w(t)=0, \qquad t<0, \end{aligned} \tag{2.1}$$

where both the order (p_0, q_0, r_0) and the matrix of coefficients

$$\Theta(p_0, q_0, r_0) = [A_1 \cdots A_{p_0} B_1 \cdots B_{q_0} C_1 \cdots C_{r_0}]^{\mathrm{T}} \tag{2.2}$$

are unknown. The assumptions required about the system (2.1) are:

A1. The noise $\{w(t)\}$ satisfies

$$\sup_t E[\|w(t)\|^{\alpha} \mid \mathscr{F}_{t-1}] < \infty \text{ a.s.,} \qquad \text{for some} \quad \alpha \geqslant 2. \tag{2.3}$$

A2. For $t \geqslant 1$, $u(t)$ is $\mathscr{F}_t$-measurable.

A3. The transfer matrix $C^{-1}(z) - \frac{1}{2} I$ is strictly positive real.

A4. The true order (p_0, q_0, r_0) belongs to a known finite set,

$$M = \{(p, q, r): 0 \leqslant p \leqslant p^*, 0 \leqslant q \leqslant q^*, 0 \leqslant r \leqslant r^*\}. \tag{2.4}$$

A5. The matrices A_{p_0}, B_{q_0}, and C_{r_0} are of row-full rank.

For a given order $(p, q, r) \in M$, the estimate of the matrix coefficients,

$$\hat{\Theta}(p, q, r, t) = [\hat{A}_{p,1}(t) \cdots \hat{A}_{p,p}(t)\, \hat{B}_{q,1}(t) \cdots \hat{B}_{q,q}(t)\, \hat{C}_{r,1}(t) \cdots \hat{C}_{r,r}(t)]^{\mathrm{T}}, \tag{2.5}$$

is obtained using the AML technique; i.e.,

$$\hat{\Theta}(p, q, r, t) = \left(\sum_{j=0}^{t-1} \Phi(p, q, r, j)\, \Phi^{\mathrm{T}}(p, q, r j) \right)^{-1} \sum_{j=0}^{t-1} \Phi(p, q, r, j)\, y^{\mathrm{T}}(j+1), \tag{2.6}$$

where

$$\begin{aligned} \Phi(p, q, r, t) = [&y^{\mathrm{T}}(t) \cdots y^{\mathrm{T}}(t-p+1)\, u^{\mathrm{T}}(t) \cdots u^{\mathrm{T}}(t-q+1) \\ &\cdot \hat{w}^{\mathrm{T}}(t) \cdots \hat{w}^{\mathrm{T}}(t-r+1)]^{\mathrm{T}}, \end{aligned} \tag{2.7}$$

with $\hat{w}(t+1)$ standing for the estimate of the noise $w(t+1)$ and here defined as

$$\hat{w}(t+1) = y(t+1) - \hat{\Theta}^{\mathrm{T}}(p^*, q^*, r^*, t+1)\, \Phi(p^*, q^*, r^*, t), \tag{2.8}$$

which is the a posteriori prediction error for the model in M with largest order.

At time n, we here estimate the order (p_0, q_0, r_0) as

$$(\hat{p}(n), \hat{q}(n), \hat{r}(n)) = \operatorname*{Arg\,Min}_{(p,q,r) \in M} PLSM(p, q, r, n), \tag{2.9}$$

where

$$PLSM(p, q, r, n) = \frac{1}{n}\left(\sum_{t=0}^{n-1} \|e(p, q, r, t+1)\|^2 + (p+q+r)\, a(n)\right)$$
$$\triangleq PLS(p, q, r, n) + (p+q+r)\, a(n)/n, \tag{2.10}$$

with

$$e(p, q, r, t+1) = y(t+1) - \hat{\theta}^{\mathrm{T}}(p, q, r, t+1)\, \Phi(p, q, r, t) \tag{2.11}$$

and $a(n)$ being a sequence of positive numbers such that

$$\left(\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right)$$
$$\times \left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right)^{c_1\delta(\alpha-2)} \Big/ a(n)$$
$$\to 0 \text{ a.s. as } n \to \infty, \qquad \text{for some} \quad c_1 > 1 \tag{2.12}$$

and

$$a(n) \Big/ \lambda_{\min}\left(\sum_{t=0}^{n-1} \Phi^0(p, q, r, t)\, \Phi^{0\mathrm{T}}(p, q, r, t)\right)$$
$$\to 0 \text{ a.s. as } n \to \infty,$$
$$\forall (p, q, r) \in \{(p_0, q^*, r^*), (p^*, q_0, r^*), (p^*, q^*, r_0)\}, \tag{2.13}$$

where δ is the Dirac measure and

$$\Phi^0(p, q, r, t) = [y^{\mathrm{T}}(t) \cdots y^{\mathrm{T}}(t-p+1)\, u^{\mathrm{T}}(t) \cdots u^{\mathrm{T}}(t-q+1)$$
$$\cdot\, w^{\mathrm{T}}(t) \cdots w^{\mathrm{T}}(t-r+1)]^{\mathrm{T}}. \tag{2.14}$$

In our set-up, we also require that

$$\Phi^{\mathrm{T}}(p, q, r, t)\, V(p, q, r, t)\, \Phi(p, q, r, t) \to 0 \text{ a.s. as } t \to \infty, \qquad \forall (p, q, r) \in M, \tag{2.15}$$

where

$$V(p, q, r, t) = \left(\sum_{j=0}^{t} \Phi(p, q, r, j)\, \Phi^{\mathrm{T}}(p, q, r, j)\right)^{-1},$$
$$V(p, q, r, -1) = cI(p, q, r),\ c \in R^+. \tag{2.16}$$

3. Main Results

We start by showing that conditions A1–A5 are sufficient for ensuring the strong consistency of the order estimates (2.9).

THEOREM 3.1. *Suppose that in the control system* (2.1) A1–A5 *hold. Then the order estimate given by* (2.9) *is strongly consistent, i.e.,*

$$(\hat{p}(n), \hat{q}(n), \hat{r}(n)) \rightarrow (p_0, q_0, r_0) \; a.s. \qquad as \quad n \rightarrow \infty. \tag{3.1}$$

Proof. We start by considering the overmodelled case, i.e., when simultaneously $p \geqslant p_0$, $q \geqslant q_0$, and $r \geqslant r_0$. For any such triple, we define

$$\Theta(p, q, r) = [A_1 \cdots A_{p_0} 0 \cdots 0 \; B_1 \cdots B_{q_0} 0 \cdots 0 \; C_1 \cdots C_{r_0} 0 \cdots 0] \tag{3.2}$$

and then rewrite (2.1) as

$$\begin{aligned} y(t+1) = \Theta^{\mathrm{T}}(p, q, r)\, \Phi^0(p, q, r, t) + w(t+1), \\ \forall p \geqslant p_0 \wedge q \geqslant q_0 \wedge r \geqslant r_0. \end{aligned} \tag{3.3}$$

We also define

$$\begin{aligned} \zeta(p, q, r, t+1) = e(p, q, r, t+1) - w(t+1), \\ \eta(t+1) = \hat{w}(t+1) - w(t+1) \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} \Phi^\eta(p, q, r, t) &= \Phi(p, q, r, t) - \Phi^0(p, q, r, t) \\ &= [0 \cdots 0 \; \eta^{\mathrm{T}}(t) \cdots \eta^{\mathrm{T}}(t-r+1)]^{\mathrm{T}}, \end{aligned} \tag{3.5}$$

and then rewrite (3.3) as

$$\begin{aligned} y(t+1) = \Theta^{\mathrm{T}}(p, q, r)\, \Phi(p, q, r, t) \\ - \Theta^{\mathrm{T}}(p, q, r)\, \Phi^\eta(p, q, r, t) + w(t+1), \end{aligned} \tag{3.6}$$

which from (2.11) implies

$$\begin{aligned} e(p, q, r, t+1) = (\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}}\, \Phi(p, q, r, t) \\ - \Theta^{\mathrm{T}}(p, q, r)\, \Phi^\eta(p, q, r, t) + w(t+1). \end{aligned} \tag{3.7}$$

Therefore, from (2.10) and (3.7)

$$\begin{aligned}
&n\,PLS(p,q,r,n)\\
&\quad=\sum_{t=0}^{n-1}\|(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))^{\mathrm{T}}\,\Phi(p,q,r,t)\|^2\\
&\qquad+2\left(\sum_{t=0}^{n-1}\Phi^{\mathrm{T}}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))\right.\\
&\qquad\left.\cdot(-\Theta^{\mathrm{T}}(p,q,r)\,\Phi^{\eta}(p,q,r,t)+w(t+1))\right)\\
&\qquad+\sum_{t=0}^{n-1}\|w(t+1)-\Theta^{\mathrm{T}}(p,q,r)\,\Phi^{\eta}(p,q,r,t)\|^2,
\end{aligned}\tag{3.8}$$

where, from the definition of $\Theta(p,q,r)$ and $\eta(t+1)$, the last term will be the same for any triple (p,q,r).

In what follows we characterize the order of the first two terms on the right hand side of (3.8).

From (2.6) and (3.2), we have

$$\begin{aligned}
&\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1)\\
&\quad=\Theta(p,q,r)-\hat{\Theta}(p,q,r,t)-V(p,q,r,t)\,\Phi(p,q,r,t)\,e_a^{\mathrm{T}}(p,q,r,t+1)
\end{aligned}\tag{3.9}$$

where

$$e_a(p,q,r,t+1)=y(t+1)-\hat{\Theta}^{\mathrm{T}}(p,q,r,t)\,\Phi(p,q,r,t).\tag{3.10}$$

Defining also

$$\begin{aligned}
\bar{Q}(p,q,r,t+1)&=(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))^{\mathrm{T}}\\
&\quad\cdot V^{-1}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))
\end{aligned}\tag{3.11}$$

and using Eq. (2.6), (2.16), the matrix inversion lemma, and Eqs. (3.4), (3.9), (3.10), and (3.11), we can show that

$$\begin{aligned}
&\operatorname{tr}(\bar{Q}(p,q,r,t+1))\\
&\quad=\operatorname{tr}(\bar{Q}(p,q,r,t))-2\Phi^{\mathrm{T}}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))\\
&\qquad\cdot\{\xi(p,q,r,t+1)-\tfrac{1}{2}(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))^{\mathrm{T}}\,\Phi(p,q,r,t)\}\\
&\qquad+2\{(e_a(p,q,r,t+1)-w(t+1))^{\mathrm{T}}\,\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\\
&\qquad-\Phi^{\mathrm{T}}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t))\}\,w(t+1)\\
&\qquad-\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\\
&\qquad\cdot(1-\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t))\,\|e_a(p,q,r,t+1)\|^2\\
&\qquad+2\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\,\|w(t+1)\|^2.
\end{aligned}\tag{3.12}$$

Summing now from $t=0$ to $t=n-1$, we get

$$\begin{aligned}
&\operatorname{tr}(\bar{Q}(p,q,r,n))-\operatorname{tr}(\bar{Q}(p,q,r,0))\\
&\quad+2\sum_{t=0}^{n-1}\Phi^{\mathrm{T}}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))\\
&\quad\cdot\{\xi(p,q,r,t+1)-\tfrac{1}{2}(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))^{\mathrm{T}}\,\Phi(p,q,r,t)\}\\
&=2\sum_{t=0}^{n-1}\{(e_a(p,q,r,t+1)-w(t+1))^{\mathrm{T}}\,\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\\
&\quad\cdot\Phi(p,q,r,t)-\Phi^{\mathrm{T}}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t))\}\,w(t+1)\\
&\quad-\sum_{t=0}^{n-1}\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\\
&\quad\cdot(1-\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t))\,\|e_a(p,q,r,t+1)\|^2\\
&\quad+2\sum_{t=0}^{n-1}\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\,\|w(t+1)\|^2. \qquad (3.13)
\end{aligned}$$

From (2.1) and (3.10) we see that $e_a(p,q,r,t+1)-w(t+1)$ is $\mathscr{F}_t-$measurable. Then, from [4], the first term on the right hand side of (3.13) has order

$$\begin{aligned}
&\sum_{t=0}^{n-1}\{(e_a(p,q,r,t+1)-w(t+1))^{\mathrm{T}}\,\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\\
&\quad-\Phi^{\mathrm{T}}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t))\}\,w(t+1)\\
&=o\Bigg(\sum_{t=0}^{n-1}\|\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)(e_a(p,q,r,t+1)-w(t+1))\\
&\quad-(\Theta(p,q,r)-\Theta(p,q,r,t))^{\mathrm{T}}\,\Phi(p,q,r,t)\|^2\Bigg)+O(1)\text{ a.s.}, \qquad (3.14)
\end{aligned}$$

or from (3.9) and the fact that $\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\leqslant 1$ a.s.,

$$\begin{aligned}
&\sum_{t=0}^{n-1}\{(e_a(p,q,r,t+1)-w(t+1))^{\mathrm{T}}\,\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\\
&\quad-\Phi^{\mathrm{T}}(p,q,r,t)(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t))\}\,w(t+1)\\
&=o\left(\sum_{t=0}^{n-1}\|(\Theta(p,q,r)-\hat{\Theta}(p,q,r,t+1))^{\mathrm{T}}\,\Phi(p,q,r,t)\|^2\right)\\
&\quad+o\left(\sum_{t=0}^{n-1}\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\|\,w(t+1)\|^2\right)+O(1)\text{ a.s.}
\end{aligned}$$

(3.15)

Therefore, from (3.10) and considering that $\mathrm{tr}(\bar{Q}(p, q, r, 0)) = O(1)$ a.s., we can rewrite (3.13) as

$$\begin{aligned}
&\mathrm{tr}(\bar{Q}(p, q, r, n)) + 2 \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1)) \\
&\quad \cdot \{\xi(p, q, r, t+1) - \tfrac{1}{2}(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}} \Phi(p, q, r, t)\} \\
&= 2(1 + o(1)) \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)\, V(p, q, r, t)\, \Phi(p, q, r, t)\, \|w(t+1)\|^2 \\
&\quad - \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)\, V(p, q, r, t)\, \Phi(p, q, r, t) \\
&\quad \cdot (1 - \Phi^{\mathrm{T}}(p, q, r, t)\, V(p, q, r, t)\, \Phi(p, q, r, t))\, \|e_d(p, q, r, t+1)\|^2 \\
&\quad + o\left(\sum_{t=0}^{n-1} \|(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}} \Phi(p, q, r, t)\|^2 \right) + O(1) \text{ a.s.}
\end{aligned} \tag{3.16}$$

For $(p, q, r) = (p^*, q^*, r^*)$, from Eq. (2.8), (3.4), and (3.6) we have

$$\begin{aligned}
\xi(p^*, q^*, r^*, t+1) &= \eta(t+1) \\
&= (\Theta(p^*, q^*, r^*) - \hat{\Theta}(p^*, q^*, r^*, t+1))^{\mathrm{T}} \Phi(p^*, q^*, r^*, t) \\
&\quad - \Theta^{\mathrm{T}}(p^*, q^*, r^*)\, \Phi^{\eta}(p^*, q^*, r^*, t),
\end{aligned} \tag{3.17}$$

which in view of (3.2) and (3.5) can be rewritten as

$$C(z^{-1})\, \eta(t+1) = (\Theta(p^*, q^*, r^*) - \hat{\Theta}(p^*, q^*, r^*, t+1))^{\mathrm{T}} \Phi(p^*, q^*, r^*, t). \tag{3.18}$$

Recalling now that from A3 $(C^{-1}(z^{-1}) - \frac{1}{2}I)$ is strictly positive real and that the second term on the right hand side of (3.16) is nonnegative, from Eq. (3.16), with $(p, q, r) = (p^*, q^*, r^*)$, and using a standard argument (see, for instance, [10]) we can show that

$$\begin{aligned}
&\sum_{t=0}^{n-1} \eta^2(t+1) \\
&= O\left(\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p^*, q^*, r^*, t)\, V(p^*, q^*, r^*, t)\, \Phi(p^*, q^*, r^*, t)\, \|w(t+1)\|^2 \right) \text{ a.s.}
\end{aligned} \tag{3.19}$$

For $(p, q, r) \neq (p^*, q^*, r^*)$ we can not rely on the same argument, since now $\xi(p, q, r, t+1)$, from (3.4) and (3.7), does not obey an equation similar to (3.18). At any rate, for any (p, q, r), from (2.6), (3.6) and (3.11) we can write

$$\begin{aligned}
&\operatorname{tr}(\bar{Q}(p, q, r, n)) \\
&\quad = \operatorname{tr}\left(\Theta^{\mathrm{T}}(p, q, r)\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{n\mathrm{T}}(p, q, r, t)\right)^{\mathrm{T}}\right. \\
&\qquad \left. \cdot V(p, q, r, n-1)\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{n\mathrm{T}}(p, q, r, t)\right) \Theta(p, q, r)\right) \\
&\qquad - 2\operatorname{tr}\left(\Theta^{\mathrm{T}}(p, q, r)\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{n\mathrm{T}}(p, q, r, t)\right)^{\mathrm{T}}\right. \\
&\qquad \left. \cdot V(p, q, r, n-1)\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, w^{\mathrm{T}}(t+1)\right)\right) \\
&\qquad + \operatorname{tr}\left(\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, w^{\mathrm{T}}(t+1)\right)^{T} V(p, q, r, n-1)\right. \\
&\qquad \left. \cdot\left(\sum_{t=0}^{n-1} (p, q, r, t)\, w^{\mathrm{T}}(t+1)\right)\right).
\end{aligned} \tag{3.20}$$

We now note that from (3.5), (3.19), and the Schwarz inequality, the first term in (3.20) has the same order as $\sum_{t=0}^{n-1} \eta^2(t+1)$, given by (3.19). Concerning the last term [5, Lemma 2],

$$\begin{aligned}
&\operatorname{tr}\left(\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, w^{\mathrm{T}}(t+1)\right)^{\mathrm{T}} V(p, q, r, n-1)\right. \\
&\qquad \left. \times\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, w^{\mathrm{T}}(t+1)\right)\right) \\
&\quad = O\left(\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)\, V(p, q, r, t)\, \Phi(p, q, r, t)\, \|w(t+1)\|^2\right) \text{ a.s.,}
\end{aligned} \tag{3.21}$$

and then, from (3.19), (3.20), and (3.21),

$$\begin{aligned}
&\operatorname{tr}(\bar{Q}((p, q, r, n)) \\
&= O\left(\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p^*, q^*, r^*, t)\, V(p^*, q^*, r^*, t)\, \Phi(p^*, q^*, r^*, t)\, \|w(t+1)\|^2\right) \text{ a.s.}
\end{aligned} \tag{3.22}$$

Now that we have characterized $\operatorname{tr}(\bar{Q}(p, q, r, n))$, we return to Eq. (3.16) and upon substituting $\xi(p, q, r, t+1)$, which from (3.4) and (3.7) is given by

$$\xi(p, q, r, t+1) = (\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}} \Phi(p, q, r, t) - \Theta^{\mathrm{T}}(p, q, r) \Phi^{\eta}(p, q, r, t), \tag{3.23}$$

we get

$$\begin{aligned} &(1+o(1)) \sum_{t=0}^{n-1} \|(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}} \Phi(p, q, r, t)\|^2 \\ &\quad + 2 \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1)) \\ &\quad \cdot (-\Theta^{\mathrm{T}}(p, q, r) \Phi^{\eta}(p, q, r, t)) \\ &= -\operatorname{tr}(\bar{Q}(p, q, r, n)) + 2(1+o(1)) \\ &\quad \cdot \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t) V(p, q, r, t) \Phi(p, q, r, t) \|w(t+1)\|^2 \\ &\quad - \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t) V(p, q, r, t) \Phi(p, q, r, t)(1 - \Phi^{\mathrm{T}}(p, q, r, t) \\ &\quad \cdot V(p, q, r, t) \Phi(p, q, r, t)) \|e_a(p, q, r, t+1)\|^2 + O(1) \text{ a.s.} \end{aligned} \tag{3.24}$$

Before substituting (3.24) into (3.8), we expand the second term on the right hand side of (3.8) [4], as

$$\begin{aligned} &\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1)) \\ &\qquad \cdot (-\Theta^{\mathrm{T}}(p, q, r) \Phi^{\eta}(p, q, r, t) + w(t+1)) \\ &\quad = \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1)) \\ &\qquad \cdot (-\Theta^{\mathrm{T}}(p, q, r) \Phi^{\eta}(p, q, r, t)) \\ &\qquad + o\left(\sum_{t=0}^{n-1} \|(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}} \Phi(p, q, r, t)\|^2\right) \\ &\qquad - (1+o(1)) \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t) V(p, q, r, t) \Phi(p, q, r, t) \|w(t+1)\|^2 \\ &\qquad + O(1) \text{ a.s.,} \end{aligned} \tag{3.25}$$

which follows from (3.10), [4], and the fact that $\Phi^{\mathrm{T}}(p, q, r, t) V(p, q, r, t) \Phi(p, q, r, t) \leqslant 1$ a.s. We can then rewrite (3.8) as

$$
\begin{aligned}
n\, PLS&(p, q, r, n) \\
&= (1+o(1)) \sum_{t=0}^{n-1} \|(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}} \Phi(p, q, r, t)\|^2 \\
&\quad + 2 \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1)) \\
&\quad \cdot (-\Theta^{\mathrm{T}}(p, q, r)\, \Phi^{\eta}(p, q, r, t)) \\
&\quad + \sum_{t=0}^{n-1} \|w(t+1) - \Theta^{\mathrm{T}}(p, q, r)\, \Phi^{\eta}(p, q, r, t)\|^2 \\
&\quad - 2(1+o(1)) \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)\, V(p, q, r, t)\, \Phi(p, q, r, t)\, \|w(t+1)\|^2 \\
&\quad + O(1) \text{ a.s.}
\end{aligned} \tag{3.26}
$$

We now note that (3.19), (3.22), (3.24), and the fact that

$$
\begin{aligned}
&\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)\, V(p, q, r, t)\, \Phi(p, q, r, t)\, \|w(t+1)\|^2 \\
&= O\left(\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p^*, q^*, r^*, t)\, V(p^*, q^*, r^*, t)\, \Phi(p^*, q^*, r^*, t)\, \|w(t+1)\|^2 \right) \text{ a.s.}
\end{aligned}
$$

imply

$$
\begin{aligned}
(1+o(1)) &\sum_{t=0}^{n-1} \|(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1))^{\mathrm{T}} \Phi(p, q, r, t)\|^2 \\
&+ 2 \sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p, q, r, t)(\Theta(p, q, r) - \hat{\Theta}(p, q, r, t+1)) \\
&\cdot (-\Theta^{\mathrm{T}}(p, q, r)\, \Phi^{\eta}(p, q, r, t)) \\
&= O\left(\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p^*, q^*, r^*, t)\, V(p^*, q^*, r^*, t)\, \Phi(p^*, q^*, r^*, t)\, \|w(t+1)\|^2 \right) \\
&+ O(1) \text{ a.s.,}
\end{aligned} \tag{3.27}
$$

and then, from (3.26) and (3.27),

$$
\begin{aligned}
n\, PLS&(p, q, r, n) \\
&= \sum_{t=0}^{n-1} \|w(t+1) - \Theta^{\mathrm{T}}(p, q, r)\, \Phi^{\eta}(p, q, r, t)\|^2 \\
&\quad + O\left(\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p^*, q^*, r^*, t)\, V(p^*, q^*, r^*, t)\, \Phi(p^*, q^*, r^*, t)\, \|w(t+1)\|^2 \right) \text{ a.s.}
\end{aligned} \tag{3.28}
$$

On the other hand, under assumption A1, from [2, Lemma 3], we have

$$\sum_{t=0}^{n-1} \Phi^{\mathrm{T}}(p^*, q^*, r^*, t)\, V(p^*, q^*, r^*, t)\, \Phi(p^*, q^*, r^*, t)\, \|w(t+1)\|^2$$
$$= O\left(\log\left(\sum_{t=0}^{n-1} \|\Phi(p^*, q^*, r^*, t)\|^2\right)\right.$$
$$\left.\cdot\left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi(p^*, q^*, r^*, t)\|^2\right)\right)^{c_2\delta(\alpha-2)}\right) \text{ a.s.,} \qquad \forall c_2 > 1, \tag{3.29}$$

and since from (3.5) and (3.19) $\sum_{t=0}^{n-1} \|\Phi(p^*, q^*, r^*, t)\|^2 = O(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2)$ a.s., from (3.28) and (3.29) it follows that

$$n\, PLS(p, q, r, n) = \sum_{t=0}^{n-1} \|w(t+1) - \Theta^{\mathrm{T}}(p, q, r)\, \Phi^{\eta}(p, q, r, t)\|^2$$
$$+ O\left(\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right.$$
$$\left.\cdot\left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right)^{c_2\delta(\alpha-2)}\right) \text{ a.s.,}$$
$$\forall c_2 > 1. \tag{3.30}$$

Finally, from (2.10), (2.12) and (3.30) with $c_2 \in (1, c_1)$, we have

$$n(PLSM(p, q, r, n) - PLSM(p_0, q_0, r_0, n))$$
$$= (p + q + r - p_0 - q_0 - r_0)\, a(n)$$
$$\cdot\left(1 + O\left(\frac{1}{a(n)} \log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right.\right.$$
$$\left.\left.\cdot\left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right)^{c_2\delta(\alpha-2)}\right)\right)$$
$$= (p + q + r - p_0 - q_0 - r_0)\, a(n)(1 + o(1)) \text{ a.s.,} \tag{3.31}$$

and since here $(p + q + r - p_0 - q_0 - r_0) > 0$, it follows that

$$PLSM(p, q, r, n) - PLSM(p_0, q_0, r_0, n) > 0 \text{ a.s.} \qquad \text{for } n \text{ large enough,} \tag{3.32}$$

and the analysis of the overmodelled case is completed.

Let us now consider the undermodelled case, i.e., $(p \vee p_0,\ q \vee q_0,\ r \vee r_0) \neq (p, q, r)$. Suppose first that $p < p_0$ and $q_0 \leqslant q \leqslant q^*$, $r_0 \leqslant r \leqslant r^*$. We start by defining

$$\begin{aligned}\tilde{\Theta}(p, q, r, t) = [&A_1 - \hat{A}_{p,1}(t) \cdots A_p - \hat{A}_{p,p}(t)\\ &B_1 - \hat{B}_{q,1}(t) \cdots B_{q0} - \hat{B}_{q,q0}(t) - \hat{B}_{q,q0+1}(t) \cdots - \hat{B}_{q,q}(t)\\ &C_1 - \hat{C}_{r,1}(t) \cdots C_{r0} - \hat{C}_{r,r0}(t) - \hat{C}_{r,r0+1}(t) \cdots - \hat{C}_{r,r}(t)]^{\mathrm{T}}\end{aligned} \tag{3.33}$$

$$\begin{aligned}\Phi(p_0, q, r, t) &= [y^{\mathrm{T}}(t) \cdots y^{\mathrm{T}}(t-p+1) \vdots y^{\mathrm{T}}(t-p) \cdots y^{\mathrm{T}}(t-p_0+1) \vdots u^{\mathrm{T}}(t) \cdots\\ &\quad u^{\mathrm{T}}(t-q+1)\, \hat{w}(t) \cdots w^{\mathrm{T}}(t-r+1)]^{\mathrm{T}}\\ &= [\Phi_y^{\mathrm{T}}(p, q, r, t) \vdots \Phi_{yd}^{\mathrm{T}}(p, q, r, t) \vdots \Phi_{u\hat{w}}^{\mathrm{T}}(p, q, r, t)]^{\mathrm{T}}\end{aligned} \tag{3.34}$$

$$\begin{aligned}\Phi(p, q, r, t) &= [y^{\mathrm{T}}(t) \cdots y^{\mathrm{T}}(t-p+1) \vdots u^{\mathrm{T}}(t) \cdots u^{\mathrm{T}}(t-q+1)\, \hat{w}(t) \cdots\\ &\quad w^{\mathrm{T}}(t-r+1)]^{\mathrm{T}}\\ &= [\Phi_y^{\mathrm{T}}(p, q, r, t) \vdots \Phi_{u\hat{w}}^{\mathrm{T}}(p, q, r, t)]^{\mathrm{T}}\end{aligned} \tag{3.35}$$

and

$$\Theta_{yd}(p, q, r) = [A_{p+1} \cdots A_{p0}]^{\mathrm{T}}. \tag{3.36}$$

We now define

$$\begin{aligned}&Q'(p, q, r, t+1)\\ &\quad = \tilde{\Theta}(p, q, r, t+1) \left(\sum_{j=0}^{t} \Phi(p, q, r, j)\, \Phi^{\mathrm{T}}(p, q, r, j) \right) \tilde{\Theta}(p, q, r, t+1)\end{aligned} \tag{3.37}$$

and recall that from (2.6) and (3.33), as in the overmodelled case we have

$$\tilde{\Theta}(p, q, r, t+1) = \tilde{\Theta}(p, q, r, t) - V(p, q, r, t)\, \Phi(p, q, r, t)\, e_a(p, q, r, t+1), \tag{3.38}$$

where $V(p, q, r, t)$ is given by (2.16) and

$$e_a(p, q, r, t+1) = y(t+1) - \hat{\Theta}(p, q, r, t)\, \Phi(p, q, r, t). \tag{3.39}$$

Considering that from (2.1), (3.5), (3.33)–(3.36), and (3.39),

$$\begin{aligned}e_a(p, q, r, t+1) = \tilde{\Theta}^{\mathrm{T}}(p, q, r, t)\, \Phi(p, q, r, t) &- \Theta_{yd}^{\mathrm{T}}(p, q, r)\, \Phi_{yd}(p, q, r, t)\\ &- \Theta^{\mathrm{T}}(p_0, q, r)\, \Phi^{\eta}(p_0, q, r, t) + w(t+1),\end{aligned} \tag{3.40}$$

we can show, from (3.37)–(3.40), that

$$
\begin{aligned}
&\sum_{t=0}^{n-1} (1-\Phi^{\mathrm{T}}(p,q,r,t)\, V(p,q,r,t)\, \Phi(p,q,r,t))\, \|e_a(p,q,r,t+1)\|^2 \\
&\quad = \operatorname{tr}\Bigg(\sum_{t=0}^{n-1} (w(t+1) + \Theta_{yd}^{\mathrm{T}}(p,q,r)\, \Phi_{yd}(p,q,r,t) \\
&\qquad - \Theta^{\mathrm{T}}(p_0,q,r)\, \Phi^{\eta}(p_0,q,r,t)) \\
&\qquad \cdot (w(t+1) + \Theta_{yd}^{\mathrm{T}}(p,q,r)\, \Phi_{yd}(p,q,r,t) - \Theta^{\mathrm{T}}(p_0,q,r)\, \Phi^{\eta}(p_0,q,r,t))^{\mathrm{T}} \Bigg) \\
&\qquad - \operatorname{tr}\left(\tilde{\Theta}^{\mathrm{T}}(p,q,r,n) \left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\, \Phi^{\mathrm{T}}(p,q,r,t) \right) \tilde{\Theta}(p,q,r,n) \right) \\
&\qquad + O(1) \text{ a.s.} \qquad (3.41)
\end{aligned}
$$

In order to analyse the first term on the right hand side of (3.41), we note that (3.5), (3.19), and (3.29) imply

$$
\begin{aligned}
&\left\| \Theta^{\mathrm{T}}(p_0,q,r) \left(\sum_{t=0}^{n-1} \Phi^{\eta}(p_0,q,r,t)\, \Phi^{\eta\mathrm{T}}(p_0,q,r,t) \right) \Theta(p_0,q,r) \right\| \\
&\quad = O\left(\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2 \right) \right. \\
&\qquad \left. \cdot \left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2 \right) \right)^{c_2\delta(\alpha-2)} \right) \text{ a.s.,} \qquad \forall c_2 > 1, \qquad (3.42)
\end{aligned}
$$

and then, from Schwarz's inequality,

$$
\begin{aligned}
&\left\| \Theta_{yd}^{\mathrm{T}}(p,q,r) \left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\, \Phi^{\eta\mathrm{T}}(p_0,q,r,t) \right) \Theta(p_0,q,r) \right\| \\
&\quad = \left\| \sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\, \Phi_{yd}^{\mathrm{T}}(p,q,r,t) \right\|^{1/2} \\
&\qquad \cdot O\left(\left(\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2 \right) \right. \right. \\
&\qquad \left. \left. \cdot \left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2 \right) \right)^{c_2\delta(\alpha-2)} \right)^{1/2} \right) \text{ a.s.,} \qquad \forall c_2 > 1. \qquad (3.43)
\end{aligned}
$$

Therefore, from (3.42), (3.43), and Chow [4], we can rewrite the first term in (3.41) as

$$\operatorname{tr}\left(\sum_{t=0}^{n-1} (w(t+1)+\Theta_{yd}^{\mathrm{T}}(p,q,r)\,\Phi_{yd}(p,q,r,t)-\Theta^{\mathrm{T}}(p_0,q,r)\,\Phi^{\eta}(p_0,q,r,t))\right.$$
$$\left.\cdot(w(t+1)+\Theta_{yd}^{\mathrm{T}}(p,q,r)\,\Phi_{yd}(p,q,r,t)-\Theta^{\mathrm{T}}(p_0,q,r)\,\Phi^{\eta}(p_0,q,r,t))^{\mathrm{T}}\right)$$
$$=\left\{1+o(1)+O\left(\frac{\begin{pmatrix}\log(\sum_{t=0}^{n-1}\|\Phi^0(p^*,q^*,r^*,t)\|^2)\\ \cdot(\log\log(\sum_{t=0}^{n-1}\|\Phi^0(p^*,q^*,r^*,t)\|^2))^{c_2\delta(\alpha-2)}\end{pmatrix}}{\lambda_{\min}(\sum_{t=0}^{n-1}\Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t))}\right)\right.$$
$$\left.+O\left(\left(\frac{\begin{pmatrix}\log(\sum_{t=0}^{n-1}\|\Phi^0(p^*,q^*,r^*,t)\|^2)\\ \cdot(\log\log(\sum_{t=0}^{n-1}\|\Phi^0(p^*,q^*,r^*,t)\|^2))^{c_2\delta(\alpha-2)}\end{pmatrix}}{\lambda_{\min}(\sum_{t=0}^{n-1}\Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t))}\right)^{1/2}\right)\right\}$$
$$\cdot\operatorname{tr}\left(\Theta_{yd}^{\mathrm{T}}(p,q,r)\left(\sum_{t=0}^{n-1}\Phi_{yd}(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r,t)\right)\Theta_{yd}(p,q,r)\right)$$
$$+\sum_{t=0}^{n-1}\|w(t+1)\|^2+O(1)\text{ a.s.,}\qquad \forall c_2>1. \tag{3.44}$$

We now turn our attention to the second term in (3.41). By defining

$$\Theta(p,q,r)=[A_1\cdots A_p\,B_1\cdots B_{q_0}\,0\cdots 0\,C_1\cdots C_{r_0}\,0\cdots 0]^{\mathrm{T}} \tag{3.45}$$

from (2.1), (3.5), and (3.35), it follows that

$$y(t+1)=\Theta^{\mathrm{T}}(p,q,r)\,\Phi(p,q,r,t)+\Theta_{yd}^{\mathrm{T}}(p,q,r)\,\Phi_{yd}(p,q,r,t)$$
$$-\Theta^{\mathrm{T}}(p_0,q,r)\,\Phi^{\eta}(p_0,q,r,t)+w(t+1) \tag{3.46}$$

and then, from (3.33) and the fact that

$$\hat{\Theta}(p,q,r,n)$$
$$=\left(\sum_{t=0}^{n-1}\Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}\left(\sum_{t=0}^{n-1}\Phi(p,q,r,t)\,y^{\mathrm{T}}(t+1)\right) \tag{3.47}$$

we get

$$\begin{aligned}\tilde{\Theta}(p, q, r, n) = &-\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)^{-1}\\ &\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi_{yd}^{\mathrm{T}}(p, q, r, t)\right) \Theta_{yd}(p, q, r)\\ &+\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)^{-1}\\ &\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\eta\mathrm{T}}(p_0, q, r, t)\right) \Theta(p_0, q, r)\\ &-\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)^{-1}\\ &\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, w^{\mathrm{T}}(t+1)\right). \end{aligned} \tag{3.48}$$

Therefore, from (3.37) and (3.48),

$$\begin{aligned}&\operatorname{tr}(Q'(p, q, r, n))\\ &\quad= \operatorname{tr}\left(\Theta_{yd}^{\mathrm{T}}(p, q, r)\left(\sum_{t=0}^{n-1} \Phi_{yd}(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)\right.\\ &\qquad\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)^{-1}\\ &\qquad\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi_{yd}^{\mathrm{T}}(p, q, r)\right) \Theta_{yd}(p, q, r)\right)\\ &\qquad+\operatorname{tr}\left(\Theta^{\mathrm{T}}(p_0, q, r)\left(\sum_{t=0}^{n-1} \Phi^{\eta}(p_0, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)\right.\\ &\qquad\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)^{-1}\\ &\qquad\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\eta\mathrm{T}}(p_0, q, r, t)\right) \Theta(p_0, q, r)\right)\\ &\qquad+\operatorname{tr}\left(\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, w^{\mathrm{T}}(t+1)\right)\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)^{-1}\right.\\ &\qquad\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p, q, r, t)\, w^{\mathrm{T}}(t+1)\right)\right)\\ &\qquad-2\operatorname{tr}\left(\Theta_{yd}^{\mathrm{T}}(p, q, r)\left(\sum_{t=0}^{n-1} \Phi_{yd}(p, q, r, t)\, \Phi^{\mathrm{T}}(p, q, r, t)\right)\right.\end{aligned}$$

$$
\begin{aligned}
&\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}\\
&\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\eta\mathrm{T}}(p_0,q,r,t)\right)\Theta(p_0,q,r)\Bigg)\\
&+2\operatorname{tr}\left(\Theta_{yd}^{\mathrm{T}}(p,q,r)\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)\right.\\
&\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,w^{\mathrm{T}}(t+1)\right)\right)\\
&-2\operatorname{tr}\left(\Theta^{\mathrm{T}}(p_0,q,r)\left(\sum_{t=0}^{n-1} \Phi^{\eta}(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)\right.\\
&\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,w^{\mathrm{T}}(t+1)\right)\right),
\end{aligned}
\tag{3.49}
$$

which, from (3.21), (3.42), and Schwarz's inequality, can be rewritten as

$$
\begin{aligned}
&\operatorname{tr}(Q'(p,q,r,n))\\
&=\operatorname{tr}\left(\Theta_{yd}^{\mathrm{T}}(p,q,r)\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)\right.\\
&\quad\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}\\
&\quad\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r)\right)\Theta_{yd}(p,q,r)\right)\\
&\quad+O\left\{\left\|\Theta_{yd}^{\mathrm{T}}(p,q,r)\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r,t)\right)^{1/2}\right\|\right.\\
&\quad\cdot\left(\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2\right)\right.\\
&\quad\left.\left.\cdot\left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2\right)\right)^{c_2\delta(\alpha-2)}\right)^{1/2}\right\}\\
&\quad+O\left(\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2\right)\right.\\
&\quad\left.\cdot\left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2\right)\right)^{c_2\delta(\alpha-2)}\right)\ \text{a.s.},\qquad \forall c_2>1.
\end{aligned}
\tag{3.50}
$$

From (3.41), (3.44) and (3.50), we then have

$$
\begin{aligned}
&\sum_{t=0}^{n-1} (1-\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t))\,\|e_a(p,q,r,t+1)\|^2 \\
&= \sum_{t=0}^{n-1} \|w(t+1)\|^2 \\
&\quad + \left\{ 1 + o(1) + O\left(\frac{\left(\begin{array}{l}\log(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2) \\ \qquad \cdot (\log\log(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2))^{c_2\delta(\alpha-2)}\end{array}\right)}{\lambda_{\min}(\sum_{t=0}^{n-1} \Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t))} \right) \right. \\
&\quad \left. + O\left(\left(\frac{\left(\begin{array}{l}\log(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2) \\ \qquad \cdot (\log\log(\sum_{t=0}^{n-1} \|\Phi^0(p^*,q^*,r^*,t)\|^2))^{c_2\delta(\alpha-2)}\end{array}\right)}{\lambda_{\min}(\sum_{t=0}^{n-1} \Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t))} \right)^{1/2} \right) \right\} \\
&\quad \cdot \operatorname{tr}\left(\Theta_{yd}^{\mathrm{T}}(p,q,r) \left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r,t) \right) \Theta_{yd}(p,q,r) \right) \\
&\quad - \operatorname{tr}\left(\Theta_{yd}^{\mathrm{T}}(p,q,r) \left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t) \right) \right. \\
&\quad \cdot \left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t) \right)^{-1} \\
&\quad \left. \cdot \left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r) \right) \Theta_{yd}(p,q,r) + O(1) \text{ a.s.}, \qquad \forall c_2 > 1, \right.
\end{aligned}
\tag{3.51}
$$

which from (2.12), (2.13) and the fact that from [5] we have, for n large enough,

$$
\begin{aligned}
&\lambda_{\min}\left(\sum_{t=0}^{n-1} \Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t) \right) \\
&\qquad \geqslant \frac{1}{3} \lambda_{\min}\left(\sum_{t=0}^{n-1} \Phi^0(p_0,q,r,t)\,\Phi^{0\mathrm{T}}(p_0,q,r,t) \right),
\end{aligned}
\tag{3.52}
$$

implies

$$\sum_{t=0}^{n-1} (1-\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t))\,\|e_a(p,q,r,t+1)\|^2$$

$$= \sum_{t=0}^{n-1} \|w(t+1)\|^2 + (1+o(1))$$

$$\times \operatorname{tr}\left\{\Theta_{yd}^{\mathrm{T}}(p,q,r)\left\{\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r,t)\right)\right.\right.$$

$$-\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}$$

$$\left.\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r,t)\right)\right\}\Theta_{yd}(p,q,r)\right\} + O(1) \text{ a.s.} \tag{3.53}$$

Let us now define

$$\overline{QF}(p,q,r,n) = \Theta_{yd}^{\mathrm{T}}(p,q,r)\left\{\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r,t)\right)\right.$$

$$-\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)$$

$$\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}$$

$$\left.\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi_{yd}^{\mathrm{T}}(p,q,r,t)\right)\right\}\Theta_{yd}(p,q,r). \tag{3.54}$$

Then, from (3.34) and (3.35) we can show that $\overline{QF}(p,q,r,n)$ can be rewritten as the quadratic form

$$\overline{QF}(p,q,r,n) = [M_1^{\mathrm{T}}(n)\;\Theta_{yd}^{\mathrm{T}}(p,q,r)\;M_2^{\mathrm{T}}(n)]$$

$$\cdot\left(\sum_{t=0}^{n-1} \Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t)\right)\begin{bmatrix} M_1(n) \\ \Theta_{yd}(p,q,r) \\ M_2(n) \end{bmatrix}, \tag{3.55}$$

where the matrices $M_1(n)$ and $M_2(n)$ are given by

$$[M_1^{\mathrm{T}}(n)\;M_2^{\mathrm{T}}(n)] = -\Theta_{yd}^{\mathrm{T}}(p,q,r)\left(\sum_{t=0}^{n-1} \Phi_{yd}(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)$$

$$\cdot\left(\sum_{t=0}^{n-1} \Phi(p,q,r,t)\,\Phi^{\mathrm{T}}(p,q,r,t)\right)^{-1}. \tag{3.56}$$

Therefore, from (3.55) and the fact that on the present case $p<p_0$, it follows that

$$\begin{aligned}\operatorname{tr}(\overline{QF}(p,q,r,n)) &\geqslant (\|M_1(n)\|^2+\|\Theta_{yd}(p,q,r)\|^2+\|M_2(n)\|^2)\\ &\quad\cdot\lambda_{\min}\left(\sum_{t=0}^{n-1}\Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t)\right)\\ &\geqslant \|\Theta_{yd}(p,q,r)\|^2\,\lambda_{\min}\left(\sum_{t=0}^{n-1}\Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t)\right)\\ &\geqslant \|A_{p0}\|^2\,\lambda_{\min}\left(\sum_{t=0}^{n-1}\Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t)\right),\end{aligned}\tag{3.57}$$

and since $\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t)\geqslant 0$ a.s., from (3.53)–(3.57) we get

$$\begin{aligned}&\sum_{t=0}^{n-1}\|e_a(p,q,r,t+1)\|^2\\ &\quad\geqslant \sum_{t=0}^{n-1}\|w(t+1)\|^2+(1+o(1))\,\|A_{p0}\|^2\\ &\qquad\cdot\lambda_{\min}\left(\sum_{t=0}^{n-1}\Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t)\right)+O(1)\text{ a.s.}\end{aligned}\tag{3.58}$$

In order to relate (3.58) with the quantity we want, namely $\sum_{t=0}^{n-1}\|e(p,q,r,t+1)\|^2$, we note that (3.7) and (3.10) imply

$$\begin{aligned}&e(p,q,r,t+1)\\ &\quad=(1-\Phi^{\mathrm{T}}(p,q,r,t)\,V(p,q,r,t)\,\Phi(p,q,r,t))\,e_a(p,q,r,t+1),\end{aligned}\tag{3.59}$$

and then from (2.15) and the fact that the right hand side of (3.58) is converging to ∞ as $n\to\infty$, we have

$$\begin{aligned}\sum_{t=0}^{n-1}\|e(p,q,r,t+1)\|^2 &\geqslant \sum_{t=0}^{n-1}\|w(t+1)\|^2+(1+o(1))\,\|A_{p0}\|^2\\ &\quad\cdot\lambda_{\min}\left(\sum_{t=0}^{n-1}\Phi(p_0,q,r,t)\,\Phi^{\mathrm{T}}(p_0,q,r,t)\right)+O(1)\text{ a.s.}\end{aligned}\tag{3.60}$$

We now recall that for the pair (p_0,q_0,r_0) from (3.13), (3.42), and Chow [4],

$$
\begin{aligned}
&n\,PLS(p_0, q_0, r_0, n)\\
&\quad= \sum_{t=0}^{n-1} \|w(t+1)\|^2 + O\left(\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right.\\
&\qquad\left.\cdot\left(\log\log\left(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2\right)\right)^{c_2\delta(\alpha-2)}\right) \text{ a.s.,} \qquad \forall c_2 > 1.
\end{aligned} \tag{3.61}
$$

On the other hand, from (3.60)

$$
\begin{aligned}
n\,PLS(p, q, r, n) \geqslant{} & \sum_{t=0}^{n-1} \|w(t+1)\|^2 + (1+o(1))\,\|A_{p_0}\|^2\\
&\cdot \lambda_{\min}\left(\sum_{t=0}^{n-1} \Phi(p_0, q, r, t)\,\Phi^{\mathrm{T}}(p_0, q, r, t)\right) + O(1) \text{ a.s.,}
\end{aligned} \tag{3.62}
$$

and then, from (2.10), (3.52), (3.61), and (3.62),

$$
\begin{aligned}
&n(PLSM(p, q, r, n) - PLSM(p_0, q_0, r_0, n))\\
&\quad\geqslant \frac{1}{3}(1+o(1))\,\|A_{p_0}\|^2\,\lambda_{\min}\left(\sum_{t=0}^{n-1} \Phi^0(p_0, q, r, t)\,\Phi^{0\mathrm{T}}(p_0, q, r, t)\right)\\
&\qquad\cdot\left\{1 + O\left(\frac{\begin{pmatrix}\log(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2)\\ \cdot(\log\log(\sum_{t=0}^{n-1} \|\Phi^0(p^*, q^*, r^*, t)\|^2))^{c_2\delta(\alpha-2)}\end{pmatrix}}{\lambda_{\min}(\sum_{t=0}^{n-1} \Phi^0(p_0, q, r, t)\,\Phi^{0\mathrm{T}}(p_0, q, r, t))}\right)\right.\\
&\qquad\left. + O\left(\frac{a(n)}{\lambda_{\min}(\sum_{t=0}^{n-1} \Phi^0(p_0, q, r, t)\,\Phi^{0\mathrm{T}}(p_0, q, r, t))}\right)\right\} + O(1) \text{ a.s.,}\\
&\qquad \forall c_2 > 1,
\end{aligned} \tag{3.63}
$$

or from (2.12) and (2.13),

$$
\begin{aligned}
&n(PLSM(p, q, r, n) - PLSM(p_0, q_0, r_0, n))\\
&\quad\geqslant \tfrac{1}{3}(1+o(1))\,\|A_{p_0}\|^2\,\lambda_{\min}\left(\sum_{t=0}^{n-1} \Phi^0(p_0, q, r, t)\,\Phi^{0\mathrm{T}}(p_0, q, r, t)\right) + O(1) \text{ a.s.}
\end{aligned} \tag{3.64}
$$

Now, from Eq. (2.12) and (2.13), for any subsequence n_s we have

$\lambda_{\min}(\sum_{t=0}^{n-1} \Phi^0(p_0, q, r, t)\, \Phi^{0\mathrm{T}}(p_0, q, r, t)) \to \infty$ a.s. as $s \to \infty$ and then, from A5 and (3.64),

$$PLSM(p, q, r, n) - PLSM(p_0, q_0, r_0, n) > 0 \text{ a.s.}$$
$$\text{for } n \text{ large enough, } p < p_0,\ q_0 \leqslant q \leqslant q^*,\ r_0 \leqslant r \leqslant r^*, \tag{3.65}$$

which implies that no triple (p, q, r) with $p < p_0$, $q_0 \leqslant q \leqslant q^*$, $r_0 \leqslant r \leqslant r^*$ can be an equilibrium point of (2.9). Arguments similar to the previous ones can be employed to show that actually (3.65) holds for any triple (p, q, r) with $(p \vee p_0, q \vee q_0, r \vee r_0) \neq (p_0, q_0, r_0)$. From this and (3.32) Theorem 3.1 is proved. ∎

Now that the strong consistency of the order estimate under conditions A1–A5 has been proved, a result similar to Guo, Chen, and Zhang [5, Theorem 2] can be established. More precisely, let $\{v(t)\}$ be a sequence of l-dimensional mutually independent random vectors independent of $\{w(t)\}$ and having the properties

$$E[v(t)] = 0,\ E[v(t)\, v^{\mathrm{T}}(t)] = \frac{I}{t^{\varepsilon}}, \qquad \|v(t)\|^2 \leqslant \frac{\sigma 2/v}{t^{\varepsilon}}, \tag{3.66}$$

with

$$\varepsilon \in \left(0, \frac{1}{2(\zeta + 1)}\right), \qquad \zeta = (m+1)\, p^* + q^* + r^* - 1. \tag{3.67}$$

Additionally, let the control $u(t)$ be generated as

$$u(t) = u^s(t) + v(t), \tag{3.68}$$

where $u^s(t)$ is an adaptive control calculated as in Chen and Guo [2], but with the addition of a second condition in the definition of σ_k, as in Hemerly and Davis [9], in order to ensure (2.15). By imposing the additional conditions:

A6. The noise $\{w(t)\}$ satisfies

$$\lim_{n \to \infty} \frac{1}{n} \sum_{t=0}^{n-1} w(t+1)\, w^{\mathrm{T}}(t+1) = R > 0 \text{ a.s.} \tag{3.69}$$

A7. $A(z^{-1})$, $B(z^{-1})$, and $C(z^{-1})$ have no common left factor,

we have

THEOREM 3.2. *Suppose the control given by* (3.68) *is applied to the system* (2.1) *and* A1, A3, A4, A6 *and* A7 *hold. If there is a nonnegative number* δ', $\delta' \in (0, (1 - 2\varepsilon(\zeta + 1))/(2\zeta + 3))$ *such that*

$$\frac{1}{n} \sum_{t=0}^{n-1} (\|y(t+1)\|^2 + \|u^s(t+1)\|^2) = O(n^{\delta'}) \text{ a.s.}, \tag{3.70}$$

then the order estimate provided by (2.9) *is strongly consistent; i.e.*,

$$(\hat{p}(n), \hat{q}(n), \hat{r}(n)) \to (p_0, q_0, r_0) \text{ a.s.} \quad \text{as} \quad n \to \infty \tag{3.71}$$

with $a(n)$, *in* (2.10), *being any sequence satisfying*

$$\frac{(\log n)(\log \log n)^{c_1 \delta(\alpha - 2)}}{a(n)} \to 0 \quad \text{as} \quad n \to \infty, \quad \text{for some} \quad c_1 > 1 \tag{3.72}$$

and

$$\frac{a(n)}{n^{1 - (\zeta + 1)(\varepsilon + \delta')}} \to 0 \quad \text{as} \quad n \to \infty. \tag{3.73}$$

Proof. The only difficulty here is to show that (2.13) holds. Considering that the proof resembles the one presented in [9], we omit the details. The reader is also referred to [5, Theorem 2]. ∎

Remark. The sequence $a(n)$ appearing in (3.72) and (3.73) is not unique. As an example we can provide $a(n) = (\log n)(\log \log n)^{\bar{c}}$, for any $\bar{c} > c_1$.

4. CONCLUSIONS

In this paper a technique for recursively estimating the order of linear stochastic system, with correlated noise and operating under feedback control, has been presented. Both the order and the coefficients estimates are strongly consistent. The consistency properties of the unmodified PLS criterion when applied to CARMA systems is a still open problem.

It is interesting to notice that in Section 2, more precisely in Eq. (2.8), the use of $\hat{w}(t+1) = y(t+1) - \hat{\Theta}^{\mathrm{T}}(p^*, q^*, r^*, t+1)\, \Phi(p^*, q^*, r^*, t)$ instead of $\hat{w}(t+1) = y(t+1) - \hat{\Theta}^{\mathrm{T}}(p, q, r, t+1)\, \Phi(p, q, r, t)$ has also a computational advantage, since the ladder form actually evaluates the prediction errors given by (2.11), $\forall (p, q, r) \in M$. See Ljung and Söderström [11, 358], for details.

References

1. H. Akaike, A new look at the statistical model identification, *IEEE Trans. Automat. Control* **14** (1974), 716–723.
2. H.-F. Chen and L. Guo, Convergence rate of least-squares identification and adaptive control for stochastic systems, *Internat. J. Control* **44** (1986), 1459–1476.
3. H.-F. Chen and L. Guo, Consistent estimation of the order of stochastic control systems, *IEEE Trans. Automat. Control* **36** (1987), 531–535.
4. Y. S. Chow, Local convergence of martingales and the law of large numbers, *Ann. Math. Statist.* **36** (1965), 552–558.
5. L. Guo, H. F. Chen and J. F. Zhang, Consistent order estimation for linear stochastic feedback control systems (CARMA model), *Automatica* **25** (1989), 147–151.
6. E. J. Hannan and B. G. Quinn, The determination of the order of an autoregression, *J. Roy. Statist. Soc. Ser. B* **41** (1979), 190–195.
7. E. J. Hannan, A. J. Mc Dougall and D. S. Poskitt, Recursive estimation of autoregressions, *J. Roy. Statist. Soc. Ser. B* **51** (1989), 217–233.
8. E. M. Hemerly and M. H. A. Davis, Strong consistency of the PLS criterion for order determination of autoregressive processes, *Ann. Statist.* **17** (1989), 941–946.
9. E. M. Hemerly and M. H. A. Davis, Recursive order estimation of stochastic control systems, *Math. Systems Theory* **22** (1989), 323–346.
10. T. L. Lai and C. Z. Wei, Extended least squares and their applications to adaptive control and prediction in linear systems, *IEEE Trans. Automat. Control* **31** (1986), 898–906.
11. L. Ljung and T. Söderström, "Theory and Practice of Recursive Identification," MIT Press, Cambridge, Mass., 1983.
12. J. Rissanen, Modelling by shortest data description, *Automatica* **14** (1978), 465–471.
13. J. Rissanen, Consistent order estimates of autoregressive processes by shortest description of data, *in* "Analysis and Optimization of Stochastic Systems" (O. Jacobs, M. Davis, M. Dempster, C. Harris, P. Parks, Eds.), pp. 451–461, Academic Press, New York, 1980.)
14. J. Rissanen, A predictive least-squares principle, *IMA J. Math. Control Inform.* **3** (1986), 211–222.
15. G. Schwarz, Estimating the dimension of a model, *Ann. Statist.* **6** (1978), 416–464.
16. S. M. Veres, "Strong Consistency of the Accumulated Prediction Error Criterion for Multivariate Processes," preprint, Imperial College, London, 1987.
17. M. Wax, Order selection for AR models by predictive least squares, *IEEE Trans. Acoust. Speech Signal Process.* **36**, No. 4 (1988).